UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/806,544	07/02/2001	Ponani Gopalakrishnan	YOR919980392US2 (473-2)	1137

75702 7590 12/28/2007

Keusey, Tutunjian & Bitetto, P.C.
20 Crossways Park North
Suite 210
Woodbury, NY 11797

EXAMINER	
BULLOCK JR, LEWIS ALEXANDER	
ART UNIT	**PAPER NUMBER**
2195	

MAIL DATE	DELIVERY MODE
12/28/2007	PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Office Action Summary	Application No.	Applicant(s)	
	09/806,544	GOPALAKRISHNAN ET AL.	
	Examiner	Art Unit	
	Lewis A. Bullock, Jr.	2195	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *10/22/07*.

2a)☒ This action is **FINAL**. 2b)☐ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *23-31,33-45,47-91 and 93-100* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *23-31,33-45,47-91 and 93-100* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☐ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☐ Information Disclosure Statement(s) (PTO/SB/08) Paper No(s)/Mail Date _____.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .
5) ☐ Notice of Informal Patent Application
6) ☐ Other: _____.

DETAILED ACTION

Claim Rejections - 35 USC § 103

1. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

2. Claims 23-31, 33-45, 47-91 and 93-100 are rejected under 35 U.S.C. 103(a) as being unpatentable over LADD (U.S. Patent 6,269,336) in view of "Multimedia Content Description in the InfoPyramid" by Li et al.

As to claim 23, LADD teaches a conversational browser (voice browser), comprising: means for interpreting a user command (voice input) and for generating a request (content request) to access a CML application (script) using CML (markup language, for example XML), wherein CML comprises meta-information (markup code implementing a conversational dialog for interaction with the user in a plurality of user interface modalities (via the network access apparatus of the system allows the user to access (i.e., view and/or hear) the information retrieved from the information source wherein the information is in the form of machine readable data, human readable data, audio or speech communications, textual information, graphical or image data, etc (col. 3, lines 40-46) (col. 3, lines 40-46; col. 4, lines 36-43; col. 4, lines 52-58); and a CML processor (parsing unit) for parsing and interpreting the meta-information (via mapping it to a tree and utilizing an interpreter to interpret the tree) to render the conversational dialog in one or more of the plurality of user interface modalities (col. 11, lines 25-49;

col. 11, line 66 – col. 12, line 24; col. 3, lines 40-46; col. 4, lines 36-43; col. 4, lines 52-58). It would be obvious to one skilled in the art that browser means for interpreting is a voice interface for receiving the voice commands. However, LADD does not explicitly mention that the CML itself comprises both GUI modality and speech modality that implements a conversational dialog to enable interaction with a user.

Li teaches a conversational markup language file or application that is accessible to a browser wherein the CML enables interaction with the user in a plurality of user interface modalities including a GUI modality and speech modality (via multimedia content is usually not in a single media format, or modality) (pg. 3789, InfoPyramids, Multi-modal) (see also wherein a news story is represented at different resolutions and is comprised of different modalities such that a user can query the database for the news story) (pg. 3792, 5.4 TV News Application). It would be obvious to one of ordinary skill in the art that the document of LADD is a news story of Li in order to retrieve and/or access the requested data / story. Therefore, it would be obvious to one of ordinary skill in the art to combine the teachings of LADD with the teachings of Li in order to facilitate the handling of multimedia the search, retrieval, manipulation, and transmission of multimedia data by providing a hierarchy for content descriptors (abstract; pg. 3789, InfoPyramids).

As to claims 24 and 25, LADD teaches a conversational browser (voice browser) of a computing device that provides a conversational user interface to render a conversational dialog (col. 11, lines 25-49). LADD also teaches that variations and

modifications may be practiced on the system (col. 2, lines 10-14). However, LADD does not teach that the browser executes on top of an operating platform. Official Notice is taken in that it is well known in the art that a browser executes on a virtual machine to send and handle remote request and therefore would be obvious in view of LADD in order to send and handle voice requests.

As to claims 26-29, LADD teaches a dialog manager (VRU server / interpreter unit) for managing and controlling the conversational dialog wherein the dialog manager allocates conversational engines (test to speech unit / automatic speech recognition unit) for rendering the conversational dialog by meta-information of a CML file (col. 9, lines 1-53; col. 13, lines 41-60).

As to claims 30, 31, 33 and 34, LADD teaches the user input command (voice input) can be input in the one or more user interface modalities (col. 11, lines 31-35; col. 3, lines 40-46; col. 4, lines 36-43; col. 4, lines 52-58; col. 2, lines 48-66), the CML is implemented in a declarative format encapsulating multi-modal dialog (col. 16, lines 5-56). Official Notice is taken in that it is well known in the art that XML is a markup language and therefore would be obvious that the markup language of LADD is XML.

As to claims 35-38, LADD teaches the input commands to the browser are voice commands (col. 11, lines 26-36). Therefore, it would be obvious to one skilled in the art that the since the commands are voice commands that navigates to a web page that the

browser implements a "what you hear is what you can say", a "say what you heard", a "say what you will hear", and a "mixed initiative dialog formats.

As to claim 80, LADD teaches a method for accessing information, comprising the steps of: processing an input command (voice input) with at least one of a plurality of conversational engines (network fetcher); generating a request (content request) based on the processed input command (voice input) to access a CML file (markup language document) from a content server (mark up language server), the CML file comprising meta-information to implement a conversational dialog in a plurality of user interface modalities (via the network access apparatus of the system allows the user to access (i.e., view and/or hear) the information retrieved from the information source wherein the information is in the form of machine readable data, human readable data, audio or speech communications, textual information, graphical or image data, etc (col. 3, lines 40-46) (col. 3, lines 40-46; col. 4, lines 36-43; col. 4, lines 52-58); transmitting the request (content request) and accessing the requested CML file from a content server using a standard networking protocol; and processing the meta-information comprising the CML file to render the conversational dialog in one or more of a plurality of user interface modalities (via parsing the information and executing the file using the browser to display and/or play sound) (col. 11, lines 25-49; col. 11, lines 66 - col. 12, line 25; col. 14, lines 3-17; col. 2, lines 20-39; col. 2, line 59 – col. 3, line 5). However, LADD does not explicitly mention that the CML itself comprises both GUI modality and

speech modality that implements a conversational dialog to enable interaction with a user.

Li teaches a conversational markup language file or application that is accessible to a browser wherein the CML enables interaction with the user in a plurality of user interface modalities including a GUI modality and speech modality (via multimedia content is usually not in a single media format, or modality) (pg. 3789, InfoPyramids, Multi-modal) (see also wherein a news story is represented at different resolutions and is comprised of different modalities such that a user can query the database for the news story) (pg. 3792, 5.4 TV News Application). It would be obvious to one of ordinary skill in the art that the document of LADD is a news story of Li in order to retrieve and/or access the requested data / story. Therefore, it would be obvious to one of ordinary skill in the art to combine the teachings of LADD with the teachings of Li in order to facilitate the handling of multimedia the search, retrieval, manipulation, and transmission of multimedia data by providing a hierarchy for content descriptors (abstract; pg. 3789, InfoPyramids).

As to claims 81 and 82, LADD teaches a conversational browser (voice browser) of a computing device executes the steps (col. 11, lines 25-49). LADD also teaches that variations and modifications may be practiced on the system (col. 2, lines 10-14). However, LADD does not teach that the browser executes on top of an operating

platform. Official Notice is taken in that it is well known in the art that a browser executes on a virtual machine to send and handle remote request and therefore would be obvious in view of LADD in order to send and handle voice requests.

As to claims 84 and 85, LADD teaches customizing the CML file (markup language document) based on the conversational capabilities of the browser (the structure of the language can be designed specifically for voice applications); and registering the capabilities with the content server (via storing the files on markup language servers) (col. 15, line 60 – col. 16, line 21).

As to claim 83, LADD teaches the steps are distributed using a conversational engine (test to speech unit / automatic speech recognition unit) and conversational arguments (request data / document attributes) (col. 11, lines 25-49; col. 9, lines 1-53; col. 13, lines 41-60).

As to claim 86-88, LADD teaches transcoding legacy content of the content server (information from the information sources) into CML based on predefined transcoding rules (via the parser unit) (col. 12, lines 15-24; col. 5, lines 8-11).

As to claim 89, LADD teaches processing the meta-information comprises playing back an audio file or generating synthesized speech output (col. 4, lines 50-61).

As to claims 90, 91 and 93, LADD teaches the CML is implemented in a declarative format encapsulating multi-modal dialog (col. 16, lines 5-56). Official Notice is taken in that it is well known in the art that XML is a markup language and therefore would be obvious that the markup language of LADD is XML (see Li reference).

As to claims 94-100, LADD teaches the CML (via markup language document) comprises one of (1) a top level element that groups other CML elements; (2) an element that specifies output to be spoken to the user (3) a menu element for encapsulating a menu that presents the user with a list of choices wherein each choice is associated with a target address identifying a CML element to visit if the corresponding choice is selected; (4) a form element for encapsulating a form that allows the user to input at least one item of information and transmit the at least one item of information to a target address; and (5) a combination thereof (col. 16, lines 29 – col. 17, line 49).

As to claim 39, LADD teaches a system for accessing information (information), comprising: a content server (mark up language server) comprising content pages (mark up language documents), wherein the content pages are implemented using a CML (mark up language) to describe a conversational dialog for interaction with a user in a plurality of user interface modalities (view and audio) including a GUI modality and speech modality (via the network access apparatus of the system allows the user to access (i.e., view and/or hear) the information retrieved from the information source

wherein the information is in the form of machine readable data, human readable data, audio or speech communications, textual information, graphical or image data, etc (col. 3, lines 40-46) (col. 15, line 60 – col. 16, line 57; col. 3, lines 40-46; col. 4, lines 36-43; col. 4, lines 52-58); and a conversational browser (voice browser) for processing a CML page received from the content server to render its conversational dialog in one or more of the plurality of user interface modalities (col. 11, lines 25-49; col. 11, line 66 – col. 12, line 24; col. 3, lines 40-46; col. 4, lines 36-43; col. 4, lines 52-58). However, LADD does not teach that the browser executes on top of an operating platform. Official Notice is taken in that it is well known in the art that a browser executes on a virtual machine to send and handle remote request and therefore would be obvious in view of LADD in order to send and handle voice requests. However, LADD does not explicitly mention that the CML itself comprises both GUI modality and speech modality that implements a conversational dialog to enable interaction with a user.

Li teaches a conversational markup language file or application that is accessible to a browser wherein the CML enables interaction with the user in a plurality of user interface modalities including a GUI modality and speech modality (via multimedia content is usually not in a single media format, or modality) (pg. 3789, InfoPyramids, Multi-modal) (see also wherein a news story is represented at different resolutions and is comprised of different modalities such that a user can query the database for the news story) (pg. 3792, 5.4 TV News Application). It would be obvious to one of ordinary skill in the art that the document of LADD is a news story of Li in order to retrieve and/or access the requested data / story. Therefore, it would be obvious to one of ordinary skill

in the art to combine the teachings of LADD with the teachings of Li in order to facilitate the handling of multimedia the search, retrieval, manipulation, and transmission of multimedia data by providing a hierarchy for content descriptors (abstract; pg. 3789, InfoPyramids).

As to claims 40-44, LADD teaches the system comprises an IVR system implemented in CML (system capable of handling a voice markup language document) (col. 11, lines 25-49; col. 14, lines 3-17) and accessibly over a packet-switched network using a standard network protocol (col. 2, lines 26-39).

As to claims 45 and 47-51, LADD teaches the CML is implemented in a declarative format encapsulating multi-modal and speech dialog (col. 16, lines 5-56; col. 16, line 58 – col. 17, line 49). Official Notice is taken in that it is well known in the art that XML is a markup language and therefore would be obvious that the markup language of LADD is XML (see Li reference).

As to claims 52-54, LADD teaches a conversational browser (voice browser) on a computing device communicating over a communications network (col. 11, lines 25-49). LADD also teaches that variations and modifications may be practiced on the system (col. 2, lines 10-14). However, LADD does not teach that the browser executes on top of an virtual machine. Official Notice is taken in that it is well known in the art that a

browser executes on a virtual machine to send and handle remote request and therefore would be obvious in view of LADD in order to send and handle voice requests.

As to claims 55 and 56, LADD teaches standard network protocols are utilized for accessing CML content pages from the content server (col. 5, lines 37-62; col. 2, lines 26-39).

As to claims 57-62, LADD teaches transcoding legacy content of the content server (information from the information sources) into CML based on predefined transcoding rules (via the parser unit) (col. 12, lines 15-24; col. 5, lines 8-11).

As to claims 63-71, LADD teaches CML (via markup language document) comprises a plurality of capability-based frames, an active link, a link to conversational data files, a link to at least one distributed conversational engine, a link to an audio file for playback, a confirmation message tag, TTS markup, scripting language and imperative code, and a link to one of a plug-in or an applet for executing a conversational task (col. 16, line 29 – col. 17, line 49).

As to claims 72-79, LADD teaches the CML (via markup language document) comprises one of (1) a top level element that groups other CML elements; (2) an element that specifies output to be spoken to the user (3) a menu element for encapsulating a menu that presents the user with a list of choices wherein each choice

is associated with a target address identifying a CML element to visit if the corresponding choice is selected; (4) a form element for encapsulating a form that allows the user to input at least one item of information and transmit the at least one item of information to a target address; and (5) a combination thereof (col. 16, lines 29 – col. 17, line 49).

Response to Arguments

3. Applicant's arguments filed October 22, 2007 have been fully considered but they are not persuasive. Applicant argues first that there is no motivation for the combination of Ladd and Li other than hindsight, by stating that Ladd is directed to a markup language to provide interactive services and Li relates to a search engine for browsing different media types and that there is a difference between the InfoPyramid of Li represented in XML and the data of different media types represented in XML. The examiner disagrees. Ladd details a voice browser that executes on a network access apparatus of the system allows the user to access (i.e., view and hear) the information retrieved from the information source (col. 3, lines 40-57). The information is provided as machine readable data, human readable data, audio or speech communications, textual information, graphical or image data, or other forms of information. The information is stored in information sources, i.e. a markup language server that includes a database, scripts and markup language documents or pages (col. 11, lines 12-19; col. 11, lines 26-36). The voice browser has a parser unit that receives the information from the network fetcher unit and parses the information according to the syntax rules of the

markup language, i.e. XML syntax) generates a tree structure and allows the interpreter unit of the voice browser to carry out a dialog with the user (col. 12, lines 7-27; col. 13, line 60). Since the voice browser allows hearing and viewing by the user, the markup language content must have visual and audio syntax that is parsed and interpreted. Li teaches that multimedia content is formatted in multiple modalities in a markup language and is available for certain query and retrieval tasks (pg. 3789, 2. Infopyramids). It is obvious based on the combination that the information in Ladd is the multimedia content of Li that is retrieved to be viewed and heard by the voice browser of Ladd since both information is retrieved based on query and retrieval tasks. Therefore, one of ordinary skill in the art would find motivation in the combination since both retrieve multimedia information in multiple modalities.

Applicant then argues that the combination does not teach a conversational browser or method for processing a CML document and rendering its conversational dialog in one or more of a plurality of user interface modalities. In essence, Applicant states that Ladd and Li teach a markup language document that may be represented in an infopyramid for describing content and fail to teach or suggest a process of parsing and interpreting CML meta-information, a CML file or CML application to render a conversational dialog of such CML file/application in one or more of a plurality of user interface modalities. The examiner disagrees. As outlined above, Ladd details a voice browser that executes on a network access apparatus of the system allows the user to access (i.e., view and hear) the information retrieved from the information source (col. 3, lines 40-57). The information is provided as machine readable data, human readable

data, audio or speech communications, textual information, graphical or image data, or other forms of information. The information is stored in information sources, i.e. a markup language server that includes a database, scripts and markup language documents or pages (col. 11, lines 12-19; col. 11, lines 26-36). The voice browser has a parser unit that receives the information from the network fetcher unit and parses the information according to the syntax rules of the markup language, i.e. XML syntax) generates a tree structure and allows the interpreter unit of the voice browser to carry out a dialog with the user (col. 12, lines 7-27; col. 13, line 60). Since the voice browser allows hearing and viewing by the user, the markup language content must have visual and audio syntax that is parsed and interpreted. Li teaches that multimedia content is formatted in multiple modalities in a markup language and is available for certain query and retrieval tasks (pg. 3789, 2. Infopyramids). Therefore, the combination teaches a browser (voice browser as outlined in Ladd) for processing a CML meta-information / CML file or CML application (scripts / information stored in a web page or web file that is formatted in a markup language as outlined in Ladd and further detailed in Li) that renders a conversational dialog in one or more user interface modalities (view and hear).

Since the references teach the language of the claims the rejection is maintained.

Conclusion

4. This is a continuation of applicant's earlier Application No. 09/806,544. All claims are drawn to the same invention claimed in the earlier application and could have been

finally rejected on the grounds and art of record in the next Office action if they had been entered in the earlier application. Accordingly, **THIS ACTION IS MADE FINAL** even though it is a first action in this case. See MPEP § 706.07(b). Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE MONTHS from the mailing date of this action. In the event a first reply is filed within TWO MONTHS of the mailing date of this final action and the advisory action is not mailed until after the end of the THREE-MONTH shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no, however, event will the statutory period for reply expire later than SIX MONTHS from the mailing date of this final action.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Lewis A. Bullock, Jr. whose telephone number is (571) 272-3759. The examiner can normally be reached on Monday-Friday, 8:30 a.m. - 5:00 p.m..

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Meng An can be reached on (571) 272-3756. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

December 21, 2007

LEWIS A. BULLOCK, JR.
PRIMARY EXAMINER